

Egan McClave

emcclave@alumni.cmu.edu | eganmcclave.com | (917) 370-3612 | github.com/eganmcclave

EDUCATION

Carnegie Mellon University - Pittsburgh, PA

M.S., Statistical Practice (GPA: 4.0)

May 2019

B.S., Mathematical Sciences - Concentration: Operations Research and Statistics

May 2018

PROFESSIONAL EXPERIENCE

DuckerFrontier – *Capstone Project Consultant*

Pittsburgh, PA / Jan-May 2019

- Completed a team-based consulting project to help DuckerFrontier effectively forecast its subscription renewal
- Used SQL to aggregate and clean 5 years of subscription data, and imputed missing values as necessary
- Applied Logistic Regression, Random Forest and various survival analysis models in R to determine key performance drivers and to predict contract renewal rates

Carnegie Mellon University – *Teaching Assistant*

Pittsburgh, PA / Sept 2017-May 2019

- Provided 4 semesters of teaching assistance in multiple advanced statistics classes with topics that include linear/nonlinear regression, predictive model-checking, uncertainty estimation, graphical models and causal inference

U.S. Army Research Lab – *Summer Intern*

Adelphi, MD / Jun-Aug 2018

- Led a 4-member team on a project that predicted network connectivity issues using Machine Learning methods
- Designed a Recurrent Neural Network using TensorFlow in Python to prevent information loss

U.S. Army Research Lab – *Summer Intern*

Adelphi, MD / Jun-Aug 2017

- Helped deploy an Internet of Things (IoT) network of multi-modal sensors to collect motion data
- Detected anomalous activities using unsupervised learning techniques in R on unstructured data
- Co-authored a research report on autonomous classifying sensor network

WebSubstance – *Summer Intern*

Sterling, VA / Jun-Aug 2016

- Assisted a digital marketing agency with web design and search engine optimization services

COURSES & PROJECTS

Data over Space and Time – *Evaluation of Wind Turbine Locations*

- Built models to analyze interdependent data over space and time
- Identified the ideal locations on Massachusetts Coast that maximize potential wind power from wind turbines with geospatial Kriging technique in R

Experimental Design & Time Series – *Analysis of London Mortality Rate*

- Studied time series models and experimental design (randomization, multiple hypothesis testing, etc...)
- Utilized Time Series Regression, Vector Autoregressive and Neural Network models in R to analyze the relationship between the environmental factors and death rates in London in early 2000s

Statistical Computing – *Detection of Anomalous Objects*

- Studied data structures and algorithms, databases, parallelism and effective programming practices
- Implemented Isolation Forest (anomaly detection algorithm) in Python to estimate speed of vehicles from a video

Data Mining – *Classification of Genomics Data*

- Learned to apply statistical methods to discover structure and make predictions from large, complex data sets
- Utilized unsupervised and supervised learning techniques to analyze and classify large scale genomics data

TECHNICAL SKILLS

R, Python, SQL, Bash, Git, Hadoop, SAS, LaTeX, HTML5/CSS, MATLAB, Mathematica, Google Analytics, Tableau

ADDITIONAL COURSEWORK

Applied Linear Models, Numerical Methods, Probability Modeling, Statistical Methods of Epidemiology, Machine Learning